

THE CLAIMS

What is claimed is:

1. A method for increasing an error tolerance of an array of m storage units, the method comprising steps of:

storing k stripes across the array of m storage units, each stripe having a plurality of elements, each stripe forming an erasure or error correcting code having a minimum Hamming distance d , and each respective element of a stripe being stored on a different storage unit;

selecting an element in a donor stripe when a difference between a minimum distance of the donor stripe and a minimum distance of a recipient stripe is greater or equal to 2, the selected element being stored on a storage unit having no elements of the recipient stripe; and

rebuilding a lost element of the recipient stripe on the selected element.

2. The method according to claim 1, wherein the minimum Hamming distance of the recipient stripe is $d \geq 2$ before the step of selecting the element in the donor stripe.

3. The method according to claim 1, further comprising a step of indicating to the storage units storing the donor stripe that the selected element has been donated before the step of rebuilding the lost element of the recipient stripe on the selected element.

4. The method according to claim 1, wherein the array of m storage units is an anamorphic array, each stripe having $n + r$ elements in which n is the number data elements in the base array, r is the number of redundant elements in the base array, $m > n + r$, $jm = k(n + r)$, and j and k are integers.

5. The method according to claim 1, wherein the storage units are hard disk

drives.

6. The method according to claim 1, wherein the storage units are RAM storage devices.

7. The method according to claim 1, wherein the donor stripe is further selected based on a minimal performance impact on the array.

8. The method according to claim 1, further comprising a step of selecting the recipient information based on an improved performance of the array.

9. The method according to claim 1, wherein the erasure or error correcting code is a parity code.

10. The method according to claim 1, wherein the erasure or error correcting code is a Winograd code.

11. The method according to claim 1, wherein the erasure or error correcting code is a symmetric code.

12. The method according to claim 1, wherein the erasure or error correcting code is a Reed-Solomon code.

13. The method according to claim 1, wherein the erasure or error correcting code is an EVENODD code.

14. The method according to claim 1, wherein the erasure or error correcting code is a derivative of an EVENODD code.

15. The method according to claim 1, wherein the array includes redundancy based on a product of a plurality of erasure or error correcting codes.

16. The method according to claim 15, wherein at least one of the erasure or error correcting codes is a parity code.

17. The method according to claim 15, wherein at least one of the erasure or error correcting codes is a Winograd code.

18. The method according to claim 15, wherein at least one of the erasure or error correcting code is a symmetric code.

19. The method according to claim 15, wherein at least one of the erasure or error correcting code is a Reed-Solomon code.

20. The method according to claim 15, wherein at least one of the erasure or error correcting code is an EVENODD code.

21. The method according to claim 15, wherein at least one of the erasure or error correcting code is a derivative of an EVENODD code.

22. The method according to claim 1, wherein when an element in the donor stripe fails during the step of rebuilding at least a portion of recipient information from the recipient stripe on the selected element, the method further comprising steps of:

terminating the step of rebuilding at least a portion of recipient information from the recipient stripe on the selected element;

selecting a second donor stripe from the plurality of stripes when a difference

ARC9-2003-0014-US1

between a minimum distance of the second donor stripe and a minimum distance of the second recipient stripe is greater or equal to 2;

selecting a donor element in the second donor stripe; and

rebuilding at least a portion of lost recipient information from the recipient stripe on the selected element in the second donor stripe.

23. The method according to claim 1, wherein when a spare element becomes available, the method further comprising a step of assigning the spare element to a selected storage unit.

24. A method of increasing the failure tolerance of an array of m storage units that is vulnerable to selected patterns of failures, comprising steps of:

storing k stripes across the array of m storage units, each stripe having a plurality of elements, each stripe forming an erasure or error correcting code having a minimum Hamming distance d , and each respective element of a stripe being stored on a different storage unit;

subsequent to an element failure, selecting a recipient element;

selecting an element in a donor stripe such that a failure tolerance of the array is increased following a rebuild operation; and

rebuilding a lost element of the recipient stripe on the selected element of the donor stripe.

25. The method according to claim 24, wherein the minimum Hamming distance of the recipient stripe is $d \geq 2$ before the step of selecting the element in the donor stripe.

26. The method according to claim 24, wherein the minimum Hamming distance of the array is increased upon completion of the step of rebuilding.

27. The method according to claim 24, wherein the recipient element is selected based on a failure pattern of the array.

28. The method according to claim 24, wherein the donor element is selected based on a predetermined target pattern.

29. The method according to claim 24, further comprising a step of indicating to the storage units storing the donor stripe that the selected element has been donated before the step of rebuilding the lost element of the recipient stripe on the selected element.

30. The method according to claim 24, wherein the array of m storage units is an anamorphic array, each stripe having $n + r$ elements in which n is the number data elements in the base array, r is the number of redundant elements in the base array, $m > n + r$, $jm = k(n + r)$, and j and k are integers.

31. The method according to claim 24, wherein the storage units are hard disk drives.

32. The method according to claim 24, wherein the storage units are RAM storage devices.

33. The method according to claim 24, wherein the erasure or error correcting code is a parity code.

34. The method according to claim 24, wherein the erasure or error correcting code is a Winograd code.

35. The method according to claim 24, wherein the erasure or error correcting

code is a symmetric code.

36. The method according to claim 24, wherein the erasure or error correcting code is a Reed-Solomon code.

37. The method according to claim 24, wherein the erasure or error correcting code is an EVENODD code.

38. The method according to claim 24, wherein the erasure or error correcting code is a derivative of an EVENODD code.

39. The method according to claim 24, wherein the array of storage units includes redundancy based on a product of a plurality of erasure or error correcting codes.

40. The method according to claim 39, wherein at least one of the erasure or error correcting codes is a parity code.

41. The method according to claim 39, wherein at least one of the erasure or error correcting codes is a Winograd code.

42. The method according to claim 39, wherein at least one of the erasure or error correcting code is a symmetric code.

43. The method according to claim 39, wherein at least one of the erasure or error correcting code is a Reed-Solomon code.

44. The method according to claim 39, wherein at least one of the erasure or error correcting code is an EVENODD code.

45. The method according to claim 39, wherein at least one of the erasure or error correcting code is a derivative of an EVENODD code.

46. A method for increasing an error tolerance of a storage system having plurality of arrays of storage units, each array having m storage units, the method comprising steps of:
storing k stripes across each respective array of m storage units, each stripe having a plurality of elements, each stripe forming an error or erasure correcting code having a minimum Hamming distance d , and each respective element of a stripe being stored on a different storage unit in the array;
selecting an element in a donor stripe when a difference between a minimum distance of the donor stripe and a minimum distance of a recipient stripe is greater or equal to 2, the selected element being stored on a storage unit having no elements of the recipient stripe; and
rebuilding a lost element of the recipient stripe on the selected element.

47. The method according to claim 46, wherein the donor stripe is stored on an array that is different from the array of the recipient stripe.

48. The method according to claim 46, wherein the donor stripe is stored on the same array as the recipient stripe.

49. The method according to claim 46, wherein the minimum Hamming distance of the recipient stripe is $d \geq 2$ before the step of selecting the element in the donor stripe.

50. The method according to claim 46, further comprising a step of indicating to the storage units storing the donor stripe that the selected element has been donated before the step of rebuilding the lost element of the recipient stripe on the selected element.

51. The method according to claim 46, wherein each array of m storage units is an anamorphic array, each stripe having $n + r$ elements in which n is the number data elements in the base array, r is the number of redundant elements in the base array, $m > n + r$, $jm = k(n + r)$, and j and k are integers.

52. The method according to claim 46, wherein the storage units are hard disk drives.

53. The method according to claim 46, wherein the storage units are RAM storage devices.

54. The method according to claim 46, wherein the selected element of the donor stripe is further selected based on a minimal performance impact on the donor stripe.

55. The method according to claim 46, wherein the donor stripe is further selected based on a minimal performance impact on the storage system.

56. The method according to claim 46, further comprising a step of selecting the recipient information based on an improved performance of the recipient stripe.

57. The method according to claim 46, further comprising a step of selecting the recipient information based on an improved performance of the storage system.

58. The method according to claim 46, wherein the erasure or error correcting code is a parity code.

59. The method according to claim 46, wherein the erasure or error correcting

code is a Winograd code.

60. The method according to claim 46, wherein the erasure or error correcting code is a symmetric code.

61. The method according to claim 46, wherein the erasure or error correcting code is a Reed-Solomon code.

62. The method according to claim 46, wherein the erasure or error correcting code is an EVENODD code.

63. The method according to claim 46, wherein the erasure or error correcting code is a derivative of an EVENODD code.

64. The method according to claim 46, wherein the array includes redundancy based on a product of a plurality of erasure or error correcting codes.

65. The method according to claim 64, wherein at least one of the erasure or error correcting codes is a parity code.

66. The method according to claim 64, wherein at least one of the erasure or error correcting codes is a Winograd code.

67. The method according to claim 64, wherein at least one of the erasure or error correcting code is a symmetric code.

68. The method according to claim 64, wherein at least one of the erasure or error correcting code is a Reed-Solomon code.

69. The method according to claim 64, wherein at least one of the erasure or error correcting code is an EVENODD code.

70. The method according to claim 64, wherein at least one of the erasure or error correcting code is a derivative of an EVENODD code.

71. The method according to claim 46, wherein when an element in the donor stripe fails during the step of rebuilding at least a portion of recipient information from the recipient stripe on the selected element, the method further comprising steps of:

terminating the step of rebuilding at least a portion of recipient information from the recipient stripe on the selected element;

selecting a second donor stripe from the plurality of stripes when a difference between a minimum distance of the second donor stripe and a minimum distance of the second recipient stripe is greater or equal to 2;

selecting a donor element in the second donor stripe; and

rebuilding at least a portion of lost recipient information from the recipient stripe on the selected element in the second donor stripe.

72. The method according to claim 46, wherein when a spare element becomes available, the method further comprising a step of assigning the spare element to a selected storage unit.

73. A data storage system, comprising;

an array of m storage units, k stripes being stored across the array of m storage units, each stripe having a plurality of elements, each stripe forming an erasure or error correcting code having a minimum Hamming distance d , and each respective element of a stripe being stored on a different storage unit; and

ARC9-2003-0014-US1

a system array controller selecting an element in a donor stripe when a difference between a minimum distance of the donor stripe and a minimum distance of a recipient stripe is greater or equal to 2, the selected element being stored on a storage unit having no elements of the recipient stripe; the system array controller rebuilding a lost element of the recipient stripe on the selected element.

74. The data storage system according to claim 73, wherein the minimum Hamming distance of the recipient stripe is $d \geq 2$ before the system array controller selects the element in the donor stripe.

75. The data storage system according to claim 73, wherein the system array controller indicates to the storage units storing the donor stripe that the selected element has been donated before the lost element of the recipient stripe is rebuilt on the selected element.

76. The data storage system according to claim 73, wherein the array of m storage units is an anamorphic array, each stripe having $n + r$ elements in which n is the number data elements in the base array, r is the number of redundant elements in the base array, $m > n + r$, $jm = k(n + r)$, and j and k are integers.

77. The data storage system according to claim 73, wherein the storage units are hard disk drives.

78. The data storage system according to claim 73, wherein the storage units are RAM storage devices.

79. The data storage system according to claim 73, wherein the system array controller selects the donor stripe further based on a minimal performance impact on the array.

80. The data storage system according to claim 73, wherein the system array controller selects the recipient information based on an improved performance of the array.

81. The data storage system according to claim 73, wherein the erasure or error correcting code is a parity code.

82. The data storage system according to claim 73, wherein the erasure or error correcting code is a Winograd code.

83. The data storage system according to claim 73, wherein the erasure or error correcting code is a symmetric code.

84. The data storage system according to claim 73, wherein the erasure or error correcting code is a Reed-Solomon code.

85. The data storage system according to claim 73, wherein the erasure or error correcting code is an EVENODD code.

86. The data storage system according to claim 73, wherein the erasure or error correcting code is a derivative of an EVENODD code.

87. The data storage system according to claim 73, wherein the array includes redundancy based on a product of a plurality of erasure or error correcting codes.

88. The data storage system according to claim 87, wherein at least one of the erasure or error correcting codes is a parity code.

89. The data storage system according to claim 87, wherein at least one of the erasure or error correcting codes is a Winograd code.

90. The data storage system according to claim 87, wherein at least one of the erasure or error correcting code is a symmetric code.

91. The data storage system according to claim 87, wherein at least one of the erasure or error correcting code is a Reed-Solomon code.

92. The data storage system according to claim 87, wherein at least one of the erasure or error correcting code is an EVENODD code.

93. The data storage system according to claim 87, wherein at least one of the erasure or error correcting code is a derivative of an EVENODD code.

94. The data storage system according to claim 73, wherein when an element in the donor stripe fails as the system array controller is rebuilding at least a portion of recipient information from the recipient stripe on the selected element, the system array controller terminates rebuilding the recipient information from the recipient stripe on the selected element, selects a second donor stripe from the plurality of stripes when a difference between a minimum distance of the second donor stripe and a minimum distance of the second recipient stripe is greater or equal to 2, selects a donor element in the second donor stripe, and rebuilds at least a portion of lost recipient information from the recipient stripe on the selected element in the second donor stripe.

95. The data storage system according to claim 73, wherein when a spare element becomes available, the system array controller assigns the spare element to a selected storage unit.

96. A data storage system, comprising:
an array of m storage units, k stripes being stored across the array of m storage units, each stripe having a plurality of elements, each stripe forming an erasure or error correcting code having a minimum Hamming distance d , and each respective element of a stripe being stored on a different storage unit; and
a system array controller selecting a recipient element subsequent to an element failure and selecting an element in a donor stripe such that a failure tolerance of the array is increased following a rebuild operation, the system array controller rebuilding a lost element of the recipient stripe on the selected element of the donor stripe.
97. The data storage system according to claim 96, wherein the minimum Hamming distance of the recipient stripe is $d \geq 2$ before the system array controller selects the element in the donor stripe.
98. The data storage system according to claim 96, wherein the minimum Hamming distance of the array is increased upon completion of rebuilding the lost element of the recipient stripe on the selected element of the donor stripe.
99. The data storage system according to claim 96, wherein system array controller selects the recipient element based on a failure pattern of the array.
100. The data storage system according to claim 96, wherein system array controller selects the donor element based on a predetermined target pattern.
101. The data storage system according to claim 96, wherein the system array controller indicates to the storage units storing the donor stripe that the selected element has been donated before the lost element of the recipient stripe is rebuilt on the selected element.

102. The data storage system according to claim 96, wherein the array of m storage units is an anamorphic array, each stripe having $n + r$ elements in which n is the number data elements in the base array, r is the number of redundant elements in the base array, $m > n + r$, $jm = k(n + r)$, and j and k are integers.

103. The data storage system according to claim 96, wherein the storage units are hard disk drives.

104. The data storage system according to claim 96, wherein the storage units are RAM storage devices.

105. The data storage system according to claim 96, wherein the erasure or error correcting code is a parity code.

106. The data storage system according to claim 96, wherein the erasure or error correcting code is a Winograd code.

107. The data storage system according to claim 96, wherein the erasure or error correcting code is a symmetric code.

108. The data storage system according to claim 96, wherein the erasure or error correcting code is a Reed-Solomon code.

109. The data storage system according to claim 96, wherein the erasure or error correcting code is an EVENODD code.

110. The data storage system according to claim 96, wherein the erasure or error

correcting code is a derivative of an EVENODD code.

111. The data storage system according to claim 96, wherein the array of storage units includes redundancy based on a product of a plurality of erasure or error correcting codes.

112. The data storage system according to claim 111, wherein at least one of the erasure or error correcting codes is a parity code.

113. The data storage system according to claim 111, wherein at least one of the erasure or error correcting codes is a Winograd code.

114. The data storage system according to claim 111, wherein at least one of the erasure or error correcting code is a symmetric code.

115. The data storage system according to claim 111, wherein at least one of the erasure or error correcting code is a Reed-Solomon code.

116. The data storage system according to claim 111, wherein at least one of the erasure or error correcting code is an EVENODD code.

117. The data storage system according to claim 111, wherein at least one of the erasure or error correcting code is a derivative of an EVENODD code.

118. A data storage system, comprising:

a plurality of arrays of storage units, each array having m storage units, k stripes being stored across each respective array of m storage units, each stripe having a plurality of elements, each stripe forming an error or erasure correcting code having a

minimum Hamming distance $d = n + 1$, and each respective element of a stripe being stored on a different storage unit in the array; and

a system array controller selecting an element in a donor stripe when a difference between a minimum distance of the donor stripe and a minimum distance of a recipient stripe is greater or equal to 2, the selected element being stored on a storage unit having no elements of the recipient stripe, the system array controller rebuilding a lost element of the recipient stripe on the selected element.

119. The data storage system according to claim 118, wherein the donor stripe is stored on an array that is different from the array of the recipient stripe.

120. The data storage system according to claim 118, wherein the donor stripe is stored on the same array as the recipient stripe.

121. The data storage system according to claim 118, wherein the minimum Hamming distance of the recipient stripe is $d \geq 2$ before the step of selecting the element in the donor stripe.

122. The data storage system according to claim 118, wherein the system array controller indicates to the storage units storing the donor stripe that the selected element has been donated before the lost element of the recipient stripe is rebuilt on the selected element.

123. The data storage system according to claim 118, wherein each array of m storage units is an anamorphic array, each stripe having $n + r$ elements in which n is the number data elements in the base array, r is the number of redundant elements in the base array, $m > n + r$, $jm = k(n + r)$, and j and k are integers.

124. The data storage system according to claim 118, wherein the storage units are ARC9-2003-0014-US1

hard disk drives.

125. The data storage system according to claim 118, wherein the storage units are RAM storage devices.

126. The data storage system according to claim 118, wherein the donor stripe is further selected based on a minimal performance impact on the donor stripe.

127. The data storage system according to claim 118, wherein the selected element of the donor stripe is further selected based on a minimal performance impact on the storage system.

128. The data storage system according to claim 118, wherein the system array controller selects the recipient information based on an improved performance of the recipient stripe.

129. The data storage system according to claim 118, wherein the system array controller selects the recipient information based on an improved performance of the storage system.

130. The data storage system according to claim 118, wherein the erasure or error correcting code is a parity code.

131. The data storage system according to claim 118, wherein the erasure or error correcting code is a Winograd code.

132. The data storage system according to claim 118, wherein the erasure or error correcting code is a symmetric code.

133. The data storage system according to claim 118, wherein the erasure or error correcting code is a Reed-Solomon code.

134. The data storage system according to claim 118, wherein the erasure or error correcting code is an EVENODD code.

135. The data storage system according to claim 118, wherein the erasure or error correcting code is a derivative of an EVENODD code.

136. The data storage system according to claim 118, wherein the array includes redundancy based on a product of a plurality of erasure or error correcting codes.

137. The data storage system according to claim 136, wherein at least one of the erasure or error correcting codes is a parity code.

138. The data storage system according to claim 136, wherein at least one of the erasure or error correcting codes is a Winograd code.

139. The data storage system according to claim 136, wherein at least one of the erasure or error correcting code is a symmetric code.

140. The data storage system according to claim 136, wherein at least one of the erasure or error correcting code is a Reed-Solomon code.

141. The data storage system according to claim 136, wherein at least one of the erasure or error correcting code is an EVENODD code.

142. The data storage system according to claim 136, wherein at least one of the erasure or error correcting code is a derivative of an EVENODD code.

143. The data storage system according to claim 118, wherein when an element in the donor stripe fails as the system array controller is rebuilding recipient information from the recipient stripe on the selected element, the system array controller terminates rebuilding recipient information from the recipient stripe on the selected element, selects a second donor stripe from the plurality of stripes when a difference between a minimum distance of the second donor stripe and a minimum distance of the second recipient stripe is greater or equal to 2, selects a donor element in the second donor stripe, and rebuilds at least a portion of lost recipient information from the recipient stripe on the selected element in the second donor stripe.

144. The data storage system according to claim 118, wherein when a spare element becomes available, the system array controller assigns the spare element to a selected storage unit.